

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 12:21:43 ; Search time 331 Seconds

(without alignments)
14019.136 Million cell updates/sec

Title: US-10-082-894-1

Perfect score: 1719

Sequence: 1 gtttaattaccacgaattga.....aaatctcgtataaaaatt 1719

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03.*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.8	19.4	1386	21	Aspergillus oryzae
C	314.2	18.3	684707	24	Listeria innocua c
C	314.2	18.3	3011208	24	Listeria innocua D
C	312.6	18.2	2944528	24	Listeria monocytog
C	284.8	16.6	580073	18	Mycoplasma genital
C	271.2	15.8	1554	24	Staphylococcus epi
C	269.6	15.7	3373	22	S. epidermidis gen
C	266.2	15.5	1518	22	S. epidermidis ope

9	243	14.1	6048	18	AAV74439	Staphylococcus aur
10	231.2	13.4	65792	22	AAF28544	Genomic fragment #
11	220.2	12.9	1620	19	AAV14157	H. pylori GHPO 476
C	219.2	12.8	11769	24	ABS53606	Human MUT/NUDIX DN
13	175.6	10.2	1041	24	ABK73405	Bacillus lichenifo
14	161.2	9.4	881	24	ABK77929	Bacillus clausii g
15	125.6	7.3	1280	23	AAS85747	DNA encoding novel
16	112.2	6.5	1888	21	AAC41181	Arabidopsis thalia
17	103.4	6.0	832	24	ABQ68729	Listeria monocytog
18	103.4	6.0	4230	24	ABQ70965	Listeria monocytog
19	98.2	5.7	1674	18	AAT86245	CDNA encoding mugw
20	98.2	5.7	1960	21	AAC45770	Arabidopsis thalia
21	95.2	5.5	2094	21	AAA26936	Essential Staphylo
22	95.2	5.5	2094	22	AAS08087	Staphylococcus aur
23	95.2	5.5	2094	22	AAF91618	Staphylococcus aur
24	92.2	5.4	294	25	ABZ55893	Aspergillus oryzae
25	91.6	5.3	1963	21	AAC41752	Arabidopsis thalia
26	89.8	5.2	1671	18	AAT86243	CDNA encoding Timo
27	88.2	5.1	1668	18	AAT86244	CDNA encoding Timo
28	86.2	5.0	1683	18	AAT86246	CDNA encoding mugw
29	84	4.9	1728	21	AAC42819	Arabidopsis thalia
30	84	4.9	1728	24	ABZ14870	Arabidopsis thalia
C	81.4	4.7	484	19	AAV69419	Enrlichia sp. HGE-
C	81.4	4.7	484	21	AAT89984	Ehrlichia antigen
C	81.4	4.7	484	24	ABS69845	Ehrlichia DNA enco
C	81.4	4.7	484	24	AAS98168	DNA encoding human
35	80.8	4.7	1593	18	AAT86242	CDNA for birch pol
36	69.4	4.0	400	18	AAV75093	Staphylococcus aur
37	68.6	4.0	254	24	ABX65706	Helicobacter pylor
38	61.6	3.6	1641	24	ABQ90454	M. capsulatus gene
39	54	3.1	294	24	ABN18433	Human ORFX polynuc
40	53.4	3.1	732	24	ABQ45492	Oligonucleotide fo
C	53.4	3.1	732	24	ABQ45493	Oligonucleotide fo
C	51.8	3.0	264	24	ABK73603	Bacillus lichenifo
43	46.8	2.7	183	24	ABN77799	Human ORF2746 CDNA
44	45.2	2.6	736	21	AAC49074	Arabidopsis thalia
45	45.2	2.6	837	21	AAC37551	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAF13546

ID AAF13546 standard; CDNA; 1386 BP.

AC AAF13546;

DT 13-MAR-2001 (first entry)

DE Aspergillus oryzae EST SEQ ID NO:6069.

XX Multiple gene expression; filamentous fungal cell; EST;

KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;

KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;

KW culture condition; environmental stress; spore morphogenesis;

KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX Aspergillus oryzae.

OS Aspergillus oryzae.

XX WO200056762-A2.

XX 28-SEP-2000.

PD 22-MAR-2000; 2000WO-US07781.

XX 22-MAR-1999; 99US-0273623.

PR (NOVO) NOVO NORDISK BIOTECH INC.

PA (NOVO) NOVO NORDISK AS.

PA Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX

WPi; 2000-594572/56.

Monitoring differential expression of genes in filamentous fungal cells
uses fluorescence-labeled nucleic acids isolated from the cells and a
substrate of expressed sequence tags -

Claim 88; Page 2502-2503; 3161pp; English.

The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway, or engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organization of the microarrays based on function of the gene products to facilitate analysis of the results. AAF07478 to AAF11247 represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

SQ Sequence 1386 BP; 317 A; 389 C; 383 G; 297 T; 0 other;

Query Match 19.4%; Score 332.8; DB 21; Length 1386;
Best Local Similarity 57.7%; Pred. No. 1.2e-76;
Matches 694; Conservative 0; Mismatches 492; Indels 16; Gaps 5;

181 CTTTCATTTGGATTGCGAGAGCTTAATGGGAATCTGAAGTTGGACATTTGAATATA 240
197 CTTGCTGTCGGATTACCGGAAGGTCATGSGGTACAGTGAAGTGGCCATCTGAACAT 256
241 GGAGCTGGAAGAGTATTATATCAAGATATTGTCGAATTAATTTGCTGCTTCAACGAAC 300
257 GGTGCGGGACGTGTGCTGCGAAGACAGTGTCTGATCGACCAACCCCTCAAGAGGCG 316
301 GAGTTTGTACAAATCCTCAGATTGTCATCAGCTGAGCGTCCAAAGAGGGAGTGT 360
317 GAGTTGAACAAGTAGACAAACGTCGTTGCATCTTCAAGCGTGCCAAAGGCAATGGC 376
361 CGATTGCAATTTATTAGGACTGGTTAGCGATGGTGTGCTCCACTCTCATATTGATCATCTT 420
377 CGTCTCATCTCTCGGCTGTGTCGCGAGCGTGGTGTCCATCCACATCACCACCTT 436
421 TTTGCGTGTACGTGTCATTTAAACAATACAAAGTGGCCAAAGGTTTTCATTCACATTTT 480
437 ATCGGCTGTGAAGGTTGCGCAAGAGATGGAGATTCCCAAGGTTTTCATCCACTCTTC 496
481 GCTGATGGTCGAGATCTTCGCCCAACAAGTGGTGTATCTTGCAACATCTTCTCAA 540
497 GCGATGAGCGGTGACACCGAACCAGAGCGGTACCAAGTACATGAGCAGTCTTCGAT 556
541 TTTATTGCTTTCGGAAGATACGGAATTTGGCTACTATTACTGGAGCTTATTATGCAATG 600
557 CAGACCAAGGAATCGGCATTGGTGAATCGCCACTGTGTTGGACGCTACTGGGCCATG 616
601 GATAGGCAACAAAGATGGGAGCGTATTAAAGATGGCTTATGAGGCAATTTGAGAGGTAT 660
617 GACCGCGAACAAGCGTGGGACCGCTGCGAAATTTGCCATGAAGGGCATTTGTCGCGGAG 676
661 GCACAAAAAGCCACCGTGTGATTAAGGCTGTCGATGTTGTTAGAGAGCGATATCTCAATC 720

677 GCGAGGAGTCTGCTCCGATC-----CTGTTAAGACCATCAACGAGCGCTACGAGAAAGGAT 730
721 GAGACTGACGAATTTCTGAAAC--AAATGTTTTTCGAGCATGGCGAGTAAAGAT 777
731 GAGACCGAGGTTCTGTAAGCCCATCATCGTGGTGGGAGGAGAGCGGTCAAGGAC 790
778 GAGGATACCTCTTATTTTCTCAATTATATCGTGTGATCGTATCGTCAAAATTTGTGAATGT 837
791 GATGACACCTTTTCTTCTTCAACTACCGCTCTGACCGTGTTCGCGAAATCACCCAACG 850
838 TTGGGTCTCGAAGTTATAAGATCTTAATAGTTTCGGTTCTCACCTTAAATAATATCAG 897
851 CTCGCT--GACTAGCAGCCGAGCCCTAAAGCCGAGCTTCCCTTACCCGAACATCCAC 907
898 ATTAGTGGGATGACCCCAATACAATAAGAGTTTCCATTTCCATTCGTTATTTCCCACTGTG 957
908 ATTACCATTATGACCCAGTATAGCCGACTACACATTCCTGTTCTTCCCTCCCTCAG 967
958 ACTCATATCAATGTGCTGCTGAATGGCTTGTCTTCAAGGAGTACTCAATTTCACTGT 1017
968 CACATGGGTAACTGCTGCTGAATGGCTCAGCAAGAGGATGTTCCAGCAATGTCACTGT 1027
1018 GCGGAACTGAGAAGTATCTCATGTTACTTCTTTTAAAGTGGTGGAGAGTTCAA 1077
1028 GCGGAGCTGAGAAGTACGCTCAGCTCACTTTCTTTTAAAGTGGTGGTATTGAGAAGCAG 1087
1078 TTCCAAGATGAAGAGGTTGTATGTTCCGTCAACAAAGAGTTGCTACATATGATTTA 1137
1088 TTCGCTGGGAGGTCCGCGACATGATCCATCCCCAA---AGTTGCGACCTACGACCTG 1144
1138 AAACCCAGAAATGAATGCTGCTGAGTTGCCGAAAAAATGTTGCGAGCAAAATGAGTCAGCG 1197
1145 GACCCCAAGATGAGCGCTGAGCGCTGCGTCAAGAGATGCGCGACCGCATTTGCTGAGGCG 1204
1198 AGGCATCTTTGTTGTTATGTCATTTTGGCGCTCTGACATGTTGGAGCATACTGCTATA 1257
1205 AAGTTGCGAGTCTGTCATGAACAACCTTCGCTCTCTGACATGGTGGCGCACACTGGTAAG 1264
1258 TTTGAACCTGCGCTCAAGCATGTCAAGCTACTGACGAGGCAATTTGGAAGATATTTGAA 1317
1265 TACGAAGCTGGCATTCAGGCTGCTGCTGCTACCGAGAGCCCATCGTGTCTATGAA 1324
1318 G-CATGCCAAACTTATAATAGTCTTATGTTTACTTCCGATCATGGAATGCTGAGAA 1376
1325 GCGTTGCAAGAGCAAGGCTATGCTCTTCACTGCGGATCATGGAACCCGAGGA 1384
1377 GA 1378
1385 AA 1386

RESULT 2
AB067196/c
ID AB067196 standard; DNA; 684707 BP.

XX AB067196;

DT 29-AUG-2002 (first entry)

XX Listeria innocua contig DNA sequence #9.

DE Antibacterial; Listeria; food contamination; mutational analysis;

KW infection; ds.

OS Listeria innocua.

XX WO200228891-A2.

PD 11-APR-2002.

XX 04-OCT-2001; 2001WO-FR03061.

XX 04-OCT-2000; 2000FR-0012697.


```
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001NO-FR03061.
XX PR 04-OCT-2000; 2000FR-0012697.
XX PA (INSP ) INST PASTEUR.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Kunst F, Glaser P;
XX DR WPI; 2002-332479/37.
XX PT New genomic sequences from Listeria species, useful for detection,
XX PT treatment and prevention of infection, also related polypeptides,
XX PT antibodies and modulators
XX PS Claim 5; SEQ ID 2058; 180pp; French.
XX CC The present invention relates to nucleic acid sequences
XX CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX CC and primers for identification and/or detection of Listeria (e.g. as
XX CC contaminants in foods, or mutational analysis) and for analysis of
XX CC gene expression. Proteins encoded by the nucleic acid sequences can be
XX CC used to screen for compounds that modulate gene expression, replication
XX CC and pathogenicity of Listeria (potential therapeutic agents), also for
XX CC treating infections by Listeria, and are useful as immunogens in
XX CC anti-Listeria vaccines.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Query Match 18.3%; Score 314.2; DB 24; Length 3011208;
Best Local Similarity 53.3%; Pred. No. 1.9e-70;
Matches 767; Conservative 0; Mismatches 648; Indels 24; Gaps 4;

QY 14 AGTTTGAGATGACAAATATCAAAATGTTCAACAAAAGCTGCTGTGTAGTTATTGATG 73
DB 2576453 ATATTTGGAGGAGTAAATAAATAAATAGTAATACCTGTAGCAATATATCTCGGATG 2576394
QY 74 GATGGGGCCCTTTCCGATGAACACACACGGGAATGCAATGCTAAAGCTAAACGCCCTATTA 133
DB 2576393 GTTTTGGTAAACGTCGAGAAACAGTAGGTAATGCTAGCTCAAGCAACAAACAAAT 2576334
QY 134 TGGACAAACTTTGTTCTGGAATTTGGCA---AAATTTGGAAGCACACGGCTTCATGTTG 190
DB 2576333 TCGACCGTTTACGTGGGCTAATTTTCTCACGGGGAACCTTAAAGCTGCTGGCCCTTGATGTTG 2576274
QY 191 GATTGGCAGAGGCTTAATGGGAAATTCCTGAAGTTGGACATTTGAATATAGGAGCTGGAA 250
DB 2576273 GTCCTCCAGAGGCCAATGGGTAACTCTGAAGTTGGCCATACAAACATCGGAGCTGGAC 2576214
QY 251 GAGTTATTATCAAGATATTGTTTCGAATTAATTTGGCTGTTCAAGCAACGAGTTGTTA 310
DB 2576213 GTATTGTCFACCAAGCTTAACTCGTATTGATAAGCAATTCGAAGCGGAATCCAGAG 2576154
QY 311 CAATTCCTCAGATTGTTGATCAGCTGAGCGGTGCAAGAGAGGGAGTGCTGCAATTCGATT 370
DB 2576153 AGAACAAAGGCCCTAAACAATGCTTTTCACTCATACAAAGAAACAAATTCGGACCTACATC 2576094
QY 371 TATTAGGACTGTTAGGAGATGTTGCTCACTCTCATATTCATGCTATTTTTCGTTGA 430
DB 2576093 TTTTCGTTTACTATACAGCGCGGGTGCATAGTCAATTAATCACTTCGTTGCACTTT 2576034
QY 431 TACGTGCACTTTAAACAATTTACAAGTGCCAAAGGTTTTCATTCACCTTTTTTGTGATGTC 490
DB 2576033 TAGAAACTCGCAAGATAAAGCGGTGAAAGAGTTTATATCCATCCATTCCTTTGATGAC 2575974
QY 491 GAGATACTTCGCCACAAAGTGAGCTGTTATCTTTGAACAACCTTCTCAATTTATGCTT 550
DB 2575974
```

```
DB 2575973 GTGACGTGGCACCAACATCCTCACTAGATAATTTAGAAAACACTAGAAAAGCTATTAGTG 2575914
QY 551 CGGAAAGTACGAGAAATTTGGCTACTATTACTGGAGCTTATTATGCAATGGATAGGACA 610
DB 2575913 ACTTAAACTATGTTGCAATTTCTACTGTTCTTGGACGTTTCTACGGGATGGATCGTGATA 2575854
QY 611 AAAGATGGAGCGGTATTAAAGATGGCTTATAGAGCAATTTGTTGGAGGATTTGGACAAAAAG 670
DB 2575853 AACGCTGGGACGCTGTTGAAAAGCATACAAAGCAATCTTAAGCGCTGAAGCTGAAGAAA - 2575795
QY 671 CCACCGTTGATAAGGCTGTCGATGTTGTTAGAGAGGATATGCTCAATCTGAGACTGACG 730
DB 2575794 -----TTTGAAGATCCAATCGAACTTCTCAAGCTTCTTACGCTAATGACAAAAATGATG 2575740
QY 731 AATTTCTGAAACCAATTTGTTTTTTCGGAGGATGGGC-----GAGTAAAGATG 778
DB 2575739 AATTCGTTGTTCTCTGCTATCACTAAAGATGGCAACCTGTTGCAACAGTTAAAGACA 2575680
QY 779 ACGATACCTCTTATTTTCTCAATATTCGTCGATGCTGATGCTGCTCAAAATTTGTAATGTT 838
DB 2575679 ACGATGCACTGATTTTCTCAATTTCCGCTGCTGACCGTCGATTCACACTTTCTAATGCT 2575620
QY 839 TGGGCTCGAAGCTTATAAAGATCTTAATAGTTCGCTTCCCTCACCTCAAAAATATTCAGA 898
DB 2575619 TCACGTATAAAGATGGGATCATTTCCGACGCTGGGGCTAACCCCTTAAAAACATCAAT 2575560
QY 899 TTAGTGGGATGACCAATACATAAAGACTTTCCATTTCCATCGTTATTTCCACCTGTGA 958
DB 2575559 TCGTTACAATGACTCTTTACAAACCAAGCATTTGATGCTGAAGTTGCTTTGAGCCANTG 2575500
QY 959 CTCATACTAATGCTTGTGTAATGCTTCTCAAGGAGTTACTCAATTTCACTGCTG 1018
DB 2575499 AAATCAAAAATGTTATTCGGCGAAGTACTTTCTTAATGAAGGTTTGTGCAACTGCGTATCG 2575440
QY 1019 CGGAAACTGAGAAAGTATTCCTCATGTTTACCTTCTTCTTAATGCTGTCGAGAAAGTTCAAT 1078
DB 2575439 CTGAAACAGAAAAATATCCACAGTAACATTTCTTATGAATGCGGACGAAATGAAGAA 2575380
QY 1079 TCCAAGATGAAGCGGTTGTTATGTTCCGTCAACCAAGAAAGTTGCTACATATGATTAA 1138
DB 2575379 TCCCTGGTGAAACCGGATTTCTAATCAATTCGCCAA---AAGTAGAACATACGATTAA 2575323
QY 1139 AACCAAGAAATGATGCTGCTGAGTTGCGGAAAAAATGCTCGAGCAAAATTTGAGTCAGGCA 1198
DB 2575322 AACCTGAAATGATGCTATGATGAAGTAAACAGATGCGCTGTTTGAAGACATTTAAAAACGACA 2575263
QY 1199 GGCATCCTTTGCTTATGTCGAATTTTGGCCCTCTCGACATGCTTGGACATCTGTTGAAT 1258
DB 2575262 AACACGATGCGATCATCTTAAACTTCGCTAATCCAGATATGTTGGACATTCAGGTTATGC 2575203
QY 1259 TTGAACCTGCGCTCAAGCATGTCAAGCTACTGACGAGGCAATTTGGAAGATATTGTAAG 1318
DB 2575202 TTGACCCAACTATTAAAGCAATCGAAGCATGATGAAATCTTGGTCTGTAGTAGATC 2575143
QY 1319 CATGCCAACTTATAATTTACGTTCTTATGGTTTACTTCCGATCATCGAATGAGTAAGA 1378
DB 2575142 TTATTTTGAAGAAAGTGGCTCAGCTATATCTTGTGACCATGCTACTCTGAAACAA 2575083
QY 1379 TGATTTGCTCCGATGGTAGCAATATGTCACATACCTGCAATTTGGTCCCATTTACT 1437
DB 2575082 TGCTTACTCCAGAGGAAACCGCACACTGCGCACACTACCGTTCCAGTTCCAGTAAT 2575024

RESULT 4
AB03041/c
ID ABA03041 standard; DNA; 2944528 BP.
XX
AC ABA03041;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes EGD-e genome sequence.
XX
```

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease; ds.
XX Listeria monocytogenes.
OS W0200717335-A2.
XX 18-OCT-2001.
XX 11-APR-2001; 2001WO-FR011118.
XX 11-APR-2000; 2000FR-0004629.
XX (INSP) INST PASTEUR.
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kraft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX WPI; 2002-010914/01.
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and
XX related polypeptides -
XX Claim 1: SEQ ID No 1; 192pp; French.
XX The present sequence is the genome sequence of Listeria monocytogenes
CC EGD-e. This sequence and fragments of this sequence are useful for
CC selecting probes and primers for detecting genes in L. monocytogenes and
CC related organisms, and to study genetic polymorphisms and other genomes.
CC Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC useful for raising specific antibodies, identification of L.
CC monocytogenes and related organisms, and for biosynthesis and
CC biodegradation, especially biosynthesis of vitamin B12. This sequence and
CC proteins encoded by it are also useful for selecting compounds that
CC regulate gene expression and cell replication and modulate L.
CC monocytogenes-related diseases. In addition, this sequence and proteins
CC encoded by it are useful in pharmaceutical and vaccines compositions for
CC the treatment or prevention of infections by L. monocytogenes and related
CC organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;
Query Match 18.28; Score 312.6; DB 24; Length 2944528;
Best Local Similarity 53.2; Pred. No. 4.8e-70;
Matches 766; Conservative 0; Mismatches 649; Indels 24; Gaps 4;
QY 14 AGTTTGAGTGGCAAAATATCAAAATGTTCAACAAAAGTCTGTCTGTAGTTATTCATG 73
DB 2529205 ATTATTGAGGCAGTAAATATGATGATGATGATGATGATGATGATGATGATG 2529146
QY 74 GATGGGCGCTTCCGATGACACACACCGGAATGCAATGCTTAAAGCTTAAACCGCTATTA 133
DB 2529145 GTTTTGTAAACGTCGACAGACAGTAGTAAATGCTAGCTCAAGCAACAAACCAAA 2529086
QY 134 TGGACAACTTGTCTGGAATTTGGCA---AAATTTGGAGACACAGCTTTCATGTTG 190
DB 2529085 TCGACCTTATTTGGCTGATTTTCCTCACGGGGAACCTTAAAGCTGTGGCTTGTATGTTG 2529026
QY 191 GATTGCCAGAGGCTTAATGGGAATTTGGAATTTGGACATTTTGAATATAGGAGCTGSA 250
DB 2529025 GCCTTCAGAGGTCGAATGGGTAACTCTGAGTTGGCCATACAAACATCGGAGCTGG 2528966
QY 251 GAGTTATTTCAAGATATTGTTGGAATTAATTTGGCTGTTTCAACGAACGAGTTTGTGA 310

DB 2528965 GTATGTCTACCAAGCTTAACTCGTATTGACAAAGCTATTGAAGAAGGGAATTCCAAG 2528906
QY 311 CAATCTCTAGATTGTTGCATCAGCTGAGCGTGCAAGAGAGGAGTGCTGATGCTATT 370
DB 2528905 AGAACAAGGCGCTAAACAATGCTTTTCACTATACAAAGAAACAACACTCGACTTACATC 2528846
QY 371 TATTAGGACTGGTTAGCGATGGTGTCTCATATTGATATTGATCATCTTTTTCGCTTGA 430
DB 2528845 TTTTCGGTTTACTATATCAGACGGCGCTGATAGTACATTAATACACTCGTTCGACTTT 2528786
QY 431 TAGCTGCATTTAAACAATTAACAAGTGCNAAGGTTTTCATTCACCTTTTTTTCGCTGATGTC 490
DB 2528785 TAGAACTCGCAAGATAAAGGCGTGAACAAGTTTATATCCATGATCTCTCTGATGAC 2528726
QY 491 GAGATACCTTCCCAACAAGTGGAGCTGGTATCTTTGAACAACCTTCTCAATTTATTGCTT 550
DB 2528725 GTGACGTGGGACCAACAATCTCTAGATAATTTAGAAACACTACAAAAGCTATTAGTG 2528666
QY 551 CGGAAAAGTACGGAGAATTTGGCTACTATTACTGGACGTTATTATGCAATGGATAGGACA 610
DB 2528665 ACTTAACTATGGCGCATTTGCTACTGTTCTGACGCTTCTAGCGATGGATCGTATA 2528606
QY 611 AAAGATGGAGCGCTTAAAGATGGCTTATGAGCAATTTGTTGGAGGTATTGGACAAAAG 670
DB 2528605 AACGCTGGGAACGTTTGAAGAAAGCATACAAAGCAATCGTAAGCGTGAAGGTGAAAAA 2528547
QY 671 CCACCGTTGATAGGCTGTGATGTTGTTAGAGAGCGATATGCTCAATCTGAGACTGACG 730
DB 2528546 -----TTTGAAGATCCCAATGCAACTTGTCAAGCTTCTTATGCTAATGACAAAATGATG 2528492
QY 731 AATTCTGAAACCAATTTGTTTTCGGACGATGGC-----GAGTAAAAAGATG 778
DB 2528491 AATTCGTTGTTCTGCTATCATTTACTAAGATGGCAACCTGTTGCAACACTTAAAGACA 2528432
QY 779 ACGATACCTCTATTTTCTTCAATTCATGCTGATGCTGATGCTGCAATTTGCTGCAATGTT 838
DB 2528431 ACGATGCAATTTTCTTCAATTTCCGCTCCTGACCGTGGATTCACCTTCTTCTAATGAT 2528372
QY 839 TGGCTCTCGACGCTTAAAGATCTTAATAGTTCGTTTCTCCACCTTAAAAATATTTCAGA 898
DB 2528371 TCACTGATAAAGAAATGGGATCATTTTCGACCGTGGGCGAGATCACCCCTAAAAACATCAAT 2528312
QY 899 TTAGTGGGATGACCAATACAAATAAGAGTTTCCATTTCCATCTGTTATTCACCACTGTGA 958
DB 2528311 TTGTTTCAATGACTCTTTACACCCCAAGATTGATGCTGAAGTTGCTTTTGAGCCATTG 2528252
QY 959 CTCATATAATGCTGCTGGAATGGCTTCTCTCAAGGAGTTACTCAATTTTCACTGCTG 1018
DB 2528251 AAATGAAAAATGTTATCGGCAAGTACTTTCTAATGAAGGCTTATCGCAACTGGGTATCG 2528192
QY 1019 CGGAACCTGAGAGTATCCTCATCTTACCTTCTTTTAAATGGTGGTGGAGAACTTCAAT 1078
DB 2528191 CTGAAACAGAAAAATATCCACACCTTAACATTTCTTTATGAATGGTGGAGAAATGAAGAT 2528132
QY 1079 TCCAAGATGAAGACGCTTGTATGCTTCCGTCCACAAAAGAGTGTGCTACATATGATTTAA 1138
DB 2528131 TCCCTGTGAAAAACCGGATCTTAATCAATTCGCCAA---AAGTAGAAACATACGATTGC 2528075
QY 1139 AACCAAAATGAATGCTGCTGGAGTTGCCGAAAAAATATGGTTCGAGCAAAATTCGAGTCAGGCA 1198
DB 2528074 AACCTGAATGAGCGCATATGAAGTAACAGATGCACTGTTGTAAGACATTTAAAAACGACA 2528015
QY 1199 GGCATCTTTGGTTATGTCGAATTTTGGCTCCTCTGACATGTTGGTGGACATCTGTTAAAT 1258
DB 2528014 AACACATGGGATCATCTTAAACTTCGCAACCCAGATATGTTGGGCACTTCAGATGTC 2527955
QY 1259 TTGAACCTGCGCTCAAAAGCATGTCAGCTACTGACGAGGCAATTTGGAAAGATATTGTAAG 1318
DB 2527954 TTGAGCCACAGATTAAAGCAATTCACGCACTAGATGAAATCTTGGTCTGTAGTAC 2527895
QY 1319 CATGCCAACTTAAATTAATGCTTCTTCTTACTTCCGATTCATGGAATGCTGAGAGA 1378

Db	1022	AAAGTATCCACATGTTAACATATCTTATGAGTGGTGACCAATGAAAGAGTTTGAAGGAG	1081
Qy	1088	AAGAGCGCTGTATGGTTCCGTCACCAAAAGTGTGTACATATGATTTAAAAACCCAGAAA	1147
Db	1082	AACGTCGTGACATCATCGATCTCCAA-- --AGTAGGCGCTTATGATTTAAAACCTGAGA	1138
Qy	1148	TGAATGCTGCTGGAGTTGCCGAAAAAATGTCGAGCAAAATTGAGTCAGGCAGGATCCCTT	1207
Db	1139	TGAGTGCAATATGAAGTTAAAGATGCAATTATTAGAAGGTTAGACAAAGGTGACTTAGATT	1198
Qy	1208	TGGTTATGTCGAATTTTGCSCCTCCGTGACATGTTTGGACATACTGGTAAATTTGAACCTG	1267
Db	1199	TAATTCTACTGAACCTTTGCTAAACCAGATATGTTGGACATAGTGGTATGCTTTGAACCAA	1258
Qy	1268	CCGTCAAGCATGTCAAGCTACTGACGAGCAATTTGGAAGATATTTTGAAGCATGCCAA	1327
Db	1259	CAATTAANGCANATCGAAGCAGTAGTAGTGCTTTGGTGAAGTCGTTGACAAATATTG	1318
Qy	1328	CTTATAAATACGTTCTTATGGTTACTTCCGATCATGGAATGCTGAGAAGATGATTCCTC	1387
Db	1319	ATATGGGTGGTCATGCCATCATCTGCACGACCCGTACTCAGATCAAGTATTAACTG	1378
Qy	1388	CCGATGGTAGTGAACATPACTGCATACCTGCAATTTGGTCCCATTACTTGCTTTCCA	1447
Db	1379	ATGACGCAACCTTATGACGACACACACAACCTAATCTGTTCCAGTTATTGTGCATTAAG	1438
Qy	1448	AAACATTTGTTTTAAATCGACTCCACCTACTGGAGAT	1485
Db	1439	AAGGTGTATCAATTAAGAAGAACTGGAGTTTAGCGAT	1476

RESULT 7

AAH54514/C

AAH54514/C
ID AAH54514 standard; DNA; 3373 BP.

AAH54514:

DT 03-SEP-2001 (first entry)

DE S. epidermidis genomic polynucleotide sequence SEQ ID NO: 3878.

Staphylococcus epidermidis SR1 strain; infection: diagnosis;

KW vaccination; endocarditis; ds.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

AA
PA (GLAX)) GLAXO GROUP LTD.

PI Kimberly WJ:

WPI: 2001-316495/33.

xx PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
PT useful for vaccinating against infections, e.g. endocarditis -

PS Claim 8; Page 1526-1527; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3373 BP: 998 A: 699 C: 510 G: 1166 T: 0 other:

Query Match	15.7%	Score	269.6;	DB	22;	Length	3373;
Best Local Similarity	51.3%	Pred. No.	5.6e-60;				
Matches	758;	Conservative	0;	Mismatches	704;	Indels	16;
						Gaps	5;

Qy	12	CAAGTTTGAGATGGACAAATATCAAAATGTTCAACAAAAGTCTGTGCTCTT--GTAGTTATTG	70
Db	2344	CAAATTGTTGAAGGTCGAAATATTTATGCGAAACCAACCACTGCCTTAATCATCTTAG	2285
Qy	71	ATGGATGGGCGCTTTCGATGAAACACACGGGAATGCAATTGCTAAAGCTAAACACGCCTA	130
Db	2284	ATGGTTTCGCAATCTGTAAGTGAACAATGCGCAATGCAAGTTAAGCAGACATAAACCTA	2225
Qy	131	TTATGGCAACACTTTGTTCTTGGAATTTGGCAAA---AATTGGAAGCACACGGTCTTCGACT	187
Db	2224	ATTTTGATCGATATTTATGAAAAATATCTACACACAAATAGAAGCTAGTGGCTTTAGATG	2165
Qy	188	TTGGATTGCCACAAGGCTTAATGGGAAATTCGTGAAGTTGGACATTTGCAATATAGGACCTG	247
Db	2164	TAGCTTCTCGTAGAGGTCAAATGGGTAACTCTGAAGTAGGACATATGAATATTGGTCGAG	2105
Qy	248	GAAGAGTATTTATCAAGATATCTTCGAAATTAATTTGGCTGTTCAACGAAACGAGTTTG	307
Db	2104	GACGCATCGTATATCAAGTTTAACTCGTATTATTAATCGATTGAAGACGGAGAATTCT	2045
Qy	308	TTACAAATCCTCAGATTGTTGCATCGCTGAGCGCTGCAAGAAGGAGGAGTGGTCGATGTC	367
Db	2044	TTGATAACACTGTATTAATAAACGCTGTAAACATGTTAAAGACAAATGGCTCTCGCGTTC	1985
Qy	368	ATTTATTAGGACTGGTTAGCGATGGTGGTGCCACTCTCATATTGATGATCACTTTTTTCGGT	427
Db	1984	ATGTAATTCGGATGCTTTCTGAAGGTGGTGACACAGTCATTATGAACATCTATTTCGTA	1925
Qy	428	TGATACGTGCATTTAAACAAATTACAAGTGCCAAAGGTTTTTCATTCACTTTTTTCTGCTGATG	487
Db	1924	TTTTAGAAATTAGCTAAAGCAAGGAATAGATAAAGTATATGTCCACGCATTTTTTAGATG	1865
Qy	488	GTCGAGATACCTCCGCAACAGTGGAGCTGGTTATCTTTGAACAACACTTCTTCAATTATTG	547
Db	1864	GTCGTGATGTTGATCAAAAATCTGCTTTGAAATATATAGAGGAACCTGAAGATAAATTTA	1805
Qy	548	CTTCGGAAGAGTACGGAGAAATGCTACTATTACTGGACGCTTATTATGCAATGATAGGG	607
Db	1804	AAGAAATTAGGTGTAGGCCAATTCGCTCTGTGTTTCAGGACGTTATTATGCTATGACCGCTG	1745
Qy	608	ACAAAAGATGGAGCGGTATTAAAGATGGCTTATGAGGCAATTTGTTGGAGGTATTGACAAA	667
Db	1744	ACAAGCGTTGGATCGTGAGGAACGTCCTATATATGCTATTTCGTAACCTTTGAAGTCCCTA	1685
Qy	668	AAGCCACCGTTGATAAGGCTGTCGATGTTGTTTAGAGACGATATGCTCAATCTCAGACTG	727
Db	1684	CATTTACTTCAGCTTAAAGCAGCGCGTTGAAAGCTAAATTATAAAATGATGTGACTGATGAAT	1625
Qy	728	ACGAATTTCTCAACCAATTTCTTTTTTCGGACGATGGCGAGTAAAGATCACCATTACTC	787
Db	1624	TCGTCGAAACCGTTTATAGTTGAGGCGCAAAACGATGGT---GTACGATGGAGCCGAC	1568
Qy	788	TTATTTCTTCAAATTATCTCGTCTGATCGTATGCGCTCAAAATTTGTGAATGTTTGGGTCTCG	847

```
Db 1567 TAATCTTTTATATTTCCCTCCAGATAGACAGCTCAACTTTTCAGAAATCTTTACTAATA 1508
QY 848 ACGTTTATAAGATCTTAAATAGTTTCCTCCACCTTAAATAATTCAGATTAGTGGGA 907
Db 1507 AAGCGTTTATGGATTAA-----AGTTGAACAAGTGACAACTTTTCTAGGCTACAT 1454
QY 908 TGACCAATACAAATAAGAGTTTCCATTTCCATGTTATCCACCTGTGACTCAATAC 967
Db 1453 TCACGAAATATAATGACATGATAGTCTGAAATTTGTTGAAAAGTTGACTTAAATA 1394
QY 968 ATGTGCTTCTGAATGGCTTCTTCAAGGAGTTTACTCAATTTCTACTGTGGGAACTG 1027
Db 1393 ATACAATCGGTGAAGTTGCTCAAGATAATGGTTGAAACAATACGATCGCTGAACTG 1334
QY 1028 AGAAGTATCTCATGTACTTCTTCTTAATGGTGTGCGAGAGTTCAATTCACAGATG 1087
Db 1333 AAAAGTATCCACATGTAACATCTTTATGAGTGGGAGCAATGAGAGTTGAGGAG 1274
QY 1088 AAGAGCGTTGTATGGTTCCTGTCACCAAAAAGAGTTGCTACATATGATTTAAACCAAAA 1147
Db 1273 AACTGCTAGACTCATCGATTCTCCAA---AAGTAGCGACTTATGATTTAAACCTGAGA 1217
QY 1148 TGAATGCTCTGAGTTGCCGAAATGCTGACCAAAATGCTGAGCAAAATGCTGAGCGCATCCTT 1207
Db 1216 TGAGTCATATGAAGTTAAAGATGCAATTTATGAGAGGTTAGACAAAGTGACTTAGATT 1157
QY 1208 TGGTTATGTGCAATTTGGCCCTCTGACATGTTGGACATATCTGTAATTTGAACCTG 1267
Db 1156 TAATCTACTGAATTTGCTAACCAGATATGTTGGACATATGTTGATGCTTGAACCAA 1097
QY 1268 CCGTCAACCATGTCAGTACTGACGAGGCAATTTGGAAGATATTTGAAGCATGCCAA 1327
Db 1096 CAATTAAGCAATCGAAGAGTAGATGAGTGTCTTGGTGAAGTCTGTGACAAAATTTATG 1037
QY 1328 CTTATAATTAAGTTCTTATGTTTACTTCGATCATGGAAATGCTGAGAGATGATGCTC 1387
Db 1036 ATATGGTGGTATGCCATCATCTGTCAGACACCGTAACCTCAGATCAAGTATTAACTG 977
QY 1388 CCGATGGTATGAACATATGCACATACCTGCAATTTGGTCCCATTTACTTCTCTTCCA 1447
Db 976 ATGACGACCAACCTATGACGACACACACAACCTAATCTCTTCCAGTTATTGAACTAAG 917
QY 1448 AAACATTTGTTTTAAATGCACTCCACCTACTGGAGAT 1485
Db 916 AAGGTGTTACATTAAGAGAAACTGGACGTTTAGCGCAT 879
```

RESULT 8

AAH53413
ID AAH53413 standard; DNA; 1518 BP.

XX AC
XX AAH53413;

XX DT
XX 03-SEP-2001 (first entry)

XX DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2219.

XX DE Staphylococcus epidermidis SR1 strain; infection; diagnosis;
XX KW vaccination; endocarditis; ds.

XX OS Staphylococcus epidermidis.

XX PN WO200134809-A2.

XX PD 17-MAY-2001.

XX PF 09-NOV-2000; 2000WO-US30782.

XX PR 09-NOV-1999; 99US-0164258.

XX PA (GLAXO) GLAXO GROUP LTD.

XX

PI Kimmerly WJ;

DR WPI; 2001-316495/33.
PT P-PSDB; AAG82563.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -

PS Claim 8; Page 594-595; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis; AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 1518 BP; 524 A; 236 C; 315 G; 443 T; 0 other;

QY Query Match 15.5%; Score 266.2; DB 22; Length 1518;
Best Local Similarity 51.0%; Pred. No. 3.2e-59;
Matches 737; Conservative 0; Mismatches 593; Indels 15; Gaps 4;

QY 44 AACAAAAAGTCGTCTTGTAGTATTGTAGTGGGCGCTTTCCGATGAACACACGGGA 103
Db 8 AACAAACCACTCCCTTAATCATCTTAGATGGTTTCGCAAACTGTAAGTGAACATGGCA 67

QY 104 ATGCAATTCGTAAGCTAAACGCGCTTATATGACAACTCTTGTCTGCAATTCGCAAA 163
Db 68 ATGCAAGTTAAGCAACACATAAACCTAATTTGTATGATATTAATAATATCTTACAA 127

QY 164 ---AATTGGAACACACACGCTCTTCATGTTGGATTCGCAAGGCTTAATGGGAATCTG 220
Db 128 CACAATAGAACCTAGTGGCTTAGATGTAGTCTTCTCGAAGTCAATGGGTAACTCTG 187

QY 221 AAGTTGGACATTTGAATATAGAGCTGGAAGAGTATTATTAAGATATTTGTCGAATTA 280
Db 188 AAGTAGGACATATGAATATTTGTCGAGGACGCTATATCAAAAGTTTAACTCGTATTA 247

QY 281 ATTTGGCTGTTCAACGAAACAGAGTTGTTACAAATCTTCAGATTTGTCATCAGCTGAGC 340
Db 248 ATAAATCGATTGAAGACGGAGAAATCTTTTGAACACTGTATTAATAACGCTGTTAAAC 307

QY 341 GTGCAAGAAGGGGAGTGGTCGATTTTATAGCACTTTATAGCACTGTTAGCGATGGTGTCC 400
Db 308 ATGTTTAAAGACAATGCTCTCGCTTCATGTATTTCGGATTTGCTTTCTGATGGGTGTAC 367

QY 401 ACTCTCATTTGATCATCTTTTGGTGGTGTGATACGTGCAATTTAAACAATTTACAGTGCCAA 460
Db 368 ACAGTCATTATAAGCATCTATTTGCTATTTTGAATTTAGCTAAAAAGCAAGGAATAGATA 427

QY 461 AGGTTTTCATTCACCTTTTTCGCTGATGTCGAGATCTTCGCAACAGTGGAGCTGGTT 520
Db 428 AAGTATATGTCACGCAATTTTATAGATGGTGTGATGATGATCAAAAATTCGTTGAAAT 487

QY 521 ATCTTGAACAACCTTCTTCAATTTATTTGCTTCGAAAAAGTACGGAGAAATTTGCTACTATTA 580
Db 488 ATATAGAGAAACTGAACATAAATTTAAGAATTAGTCTAGGCCAATTCGCTTCTGTTT 547

Matches	709;	Conservative	1;	Mismatches	701;	Indels	22;	Gaps	6;
Qy	12	CAAGTTTGAGATGGACAAATATCAAAATGTTCACACAAAAGCTGTGCTT-GTAGTTATTG	70						
Db	4390	CAATTGTTAGAGGTGCAAAATATATCGCTAAGAAACCAACTGCGCTTAATTATTTTAG	4449						
Qy	71	ATGGATGGGCGCTTTCGATGAACAACACGGGAATGCAATTGCTTAAAGCTTAAACGCCTA	130						
Db	4450	ATGGTTTTCGGAACCGCGAAGCAACATGGTAAATGCGGTAATAATTAGCAACAAGCCTA	4509						
Qy	131	TTATGGACAAACTTTGTCTCTGGAAT---TGGCAAAAATTTGGAAGCACACGGCTCTTCATG	187						
Db	4510	ATTTTGATCGTTATTACAACAAATATCCAAGCACTCAATCGAAGCAGTGGCTTAGATG	4569						
Qy	188	TTGGATTGCCAAGAGGCTTAATGGGAAATTCCTGAAGTTGGACATTTGCAATATAGGAGCTG	247						
Db	4570	TTGGTACTCTGAGGACAAATGGGTAACTCAGAAGTTGTTCATATGAATATCGGTGCAG	4629						
Qy	248	GAAGAGTTATTTATCAAGATATTGTTCCGAATTAATTTGGCTGTTTCAACGAAACGAGTTTG	307						
Db	4630	GACGTATCGTTTATCAAGTTTAACTCGAATCAATAATCAATGAGACGGTGATTTCT	4689						
Qy	308	TTACAATCTCAGATTTGTTGCATCAGCTGAGCGTGCAAGAGGGGAGTGGTCGATGTC	367						
Db	4690	TTGAAAATGATGTTTTAAATAATGCAATTTGCACACGTGAATTCACATGATTCACGGTTAC	4749						
Qy	368	ATTTATTAGCACTGGTTTAGCGATGGTGGTGTCACACTCTCATATGTATGCATCTTTTTCGGT	427						
Db	4750	ACATCTTTGGTTATTGCTCTGACGGTGGTGACACAGTCATACACAATTTATTTGCTT	4809						
Qy	428	TGATACGTGCATTTAAACAATTACAAGTGCCAAAGGTTTTTCATTCATCTTTTTTGTGCTGATG	487						
Db	4810	TGTTAGAACTTGCTTAAAAACAAGGTGTTGAAAAAGTTTACGTACACGCATTTTATAGT	4869						
Qy	488	GTCCAGATACTTCGCCAACAAGTGGAGCTGGTTATCTTGAACAACACTTCTTCAATTTATG	547						
Db	4870	GCCGTGACGTAGATCAAAAAATCCGCTTTGAAATACATCGAAGAGACATGAAAGCTTAAATTC	4929						
Qy	548	CTTCGGAAAAAGTACGGGAATTTGGCTACTATTACTGACGCTTATTATGCAATGGATAGGG	607						
Db	4930	ATGNAATTAGCAATTTGGTCAATTTGCACTCTGCTGCTGCTTATATGCAATGGATCGTG	4989						
Qy	608	ACAAAAGATGGAGCGTATTAAAGATGGCTTATGAGGCAATTTGTTGGAGGTATTTGGACAAA	667						
Db	4990	ACAAACGTTGGGACGCTGAAGAAAAAGCTTACAAATGCTATT-----CGTAATTTTGATG	5043						
Qy	668	AAGCCACGTTGTATAGCGCTGTCGATCTGTTAGAGAGCGATATGCTCAATCTGAGACTG	727						
Db	5044	CCCCAATTATGCAACTGCCAAGAAGGTGTAGAACGACGCTAATATGAGGGCTTTAACTG	5103						
Qy	728	ACGAATTTCTGAACCA---ATTGTTTTTTCGGACGATGGCGGAGTAAAAAGATGACGATA	784						
Db	5104	ACGAATTCGTAGTACCAATTCATCGTTGAGAAATCAAAATGACGGTGTTAATGATGGAGNN	5163						
Qy	785	CTCTTATTCTTCAATTATCGTGCTGATCGTATGCGTCAAAATTTGGAATGTTTGGGTC	844						
Db	5164	NN	5223						
Qy	845	TCGAACGTTTATAAGATCTTAAATAGTTTCGGTCTCTCACCCCTAAAAATATTCAGATTAGT	904						
Db	5224	ACAGACGATTCGAAGGCTTTAAA-----CTTGAACAAGTTAAAGACTTATCTATGCAA	5277						
Qy	905	GGATGACCCCAATACATAAAGAGTTTCCCATTTCCATCGTTATTTCCCACTGTGACTCAT	964						
Db	5278	CATTCACTAAGTATAATGACAAATATCGATCGGGTATCGCTTTCGAAAAAGTTGATTTAA	5337						
Qy	965	CTAATGTCCTGCTGAATGCTTGTCTTCAAGGAGTTACTCAATTTCACTGTGCGGAAA	1024						
Db	5338	ATATACAAATTTGGTGAATTTGCACAAATTAACAATTTAACTCAATTCGTTATCGAGAA	5397						
Qy	1025	CTGAGAAGTATCCTCATGTTTACCTTCTTTTAAATGGTGGTCGAGAAGTTTCAATTTCCAAG	1084						
Db	5398	CTGAAAAATCCCTCAGCTTACTTTTATGAGTGGTGGACCTTAACGAGGAATTTAAAG	5457						

Qy	1085	ATGAAGACCGCTGTGATGGTTCCGTACCACAAAGAAGTGTGTACATATGA	TTTAAAACCCAG	1144
Dd	5458	GTGAACGCCCTCGTTTAATTGATTACCTA--	--AAGTTGCAACGTATGACTTGAAACCCAG	5514
Qy	1145	AAATGAATGCTGTGGAGTTGCCGAAAAAATGTCGAGCAAATGAGT	CAGGCAGGCATC	1204
Dd	5515	AAATGAGTGTCTATGAAGTTAAAGATGCA	TATTAGAAGAGTTAAATAAGGTGACTTGG	5574
Qy	1205	CTTTGGTTATGTGCAATTTTTGCGCCTCCTGACATGGTTGGACAT	TACTGGTAAATTTTGAAC	1264
Dd	5575	ACTTAATATTATTAACCTTTGCTAACCTTGATGTTGGACAT	TGGTATGCTTTCGAGC	5634
Qy	1265	CTCGCGTCAAGCATGTCAAGCTACTGACGAGGCAATTTGAAAGAT	TATTTGAAGCATGCC	1324
Dd	5635	CGACAATCAAGCAATCAAGACCGGTTGATGAATGTTTAGGAGAAG	TGGTTGATAAAGATTT	5694
Qy	1325	AAACTTATAATTACGTTCTTATGTTTCTTACTTCCGATCATG	CAAAATGCTGAGAAGATGATG	1384
Dd	5695	TAGACATGGACGGTTATGCAATTTATTACTGCTGACCATG	GTAATCTGTATCAAGTATTGA	5754
Qy	1385	CTCCCGATGGTAGTGAACATPACTGCATACCTGCAATTTGGT	CCCCATTACT	1437
Dd	5755	CGGATGATGATCAACCAATGACTACTGACGAWACACGAAC	CCCAGTACCAGTGATT	5807
 RESULT 10 AAF28544 ID AAF28544 standard; DNA; 65792 BP. XX AC AAF28544; DT DT 04-APR-2001 (first entry) XX DE Genomic fragment #31. XX KW Genomic library; bacteria; human upper airway; otitis media; sinusitis; KW bronchopulmonary; endocarditis; meningitis; ss. XX OS Moraxella catarrhalis. PN WO200078968-A2. XX PD 28-DEC-2000. XX PF 16-JUN-2000; 2000WO-US16649. XX PR 18-JUN-1999; 99US-0140121. XX PA (INCY-) INCYTE GENOMICS INC. XX PI Lagace RE, Patterson C, Berg KL; XX WPI; 2001-041427/05. XX PT Genomic library for identifying diagnostic and therapeutic PT compositions, and for identifying virulence factors, regulatory PT elements and drug targets, comprises Moraxella catarrhalis nucleic PT acids - XX PS Claim 1; Page 278-293; 545pp; English. XX CC The present invention relates to a Moraxella catarrhalis genomic library CC comprising of a combination of 41 nucleic acid molecules (see CC AAF28514-AAF28554). The library has a number of uses described in the CC specification e.g. is useful for identifying diagnostic and therapeutic CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large CC aerobic, gram-negative diplococcus, normally found among the bacterial CC flora of human upper airways. M. catarrhalis is known to cause acute, CC localised infections such as otitis media, sinusitis and bronchopulmonary CC infection and life-threatening, systemic diseases including endocarditis CC and meningitis.				

CC infections, including acute, chronic, and atrophic gastritis, and peptic
CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC for the production of antibodies. The products can also be used for
CC detection and diagnosis.

XX
SQ Sequence 1620 BP; 518 A; 324 C; 348 G; 430 T; 0 other;

Query Match 12.9%; Score 222.2; DB 19; Length 1620;
Best Local Similarity 53.9%; Pred. No. 1e-47;
Matches 655; Conservative 0; Mismatches 518; Indels 42; Gaps 8;

QY 167 TGAAGCAGACGGCTCTTCATGTTGGATGCGCAGAGGCTTAATGGAATTCGAAAGTTG 226
DB 196 TTGATACGATGCTTTCAGCGTGGCTTACCTAAGGGGCAAAATGGAATTCGAAAGTTG 255
QY 227 GACATTTGAATATGAGCGTGGAGAGTTATTTATCAAGATATGTTCCGATTAATTTGG 286
DB 256 GGCATATGTTATGCGGCTGGTAGGGTCTCTATCAGGATTTAGTCAAAATTTCTTTAA 315
QY 287 CTGTTCAAGCAACAGGTTTGTACAAATCCTCAGATTTGTCATCAGCTGAGCGTCAAA 346
DB 316 GCCTTCAACAGATGAATTAACAAACACCCCGCTTTT-----AACACGATCCAA 368
QY 347 AGAAGGGAGTGTGCGATTCATTTATAGGACTGTTAGCGATGTTGGTTCACCTCTC 406
DB 369 AAAAGCCCTGTG-----TGCATCTTATGGTTTAATGAGCGATGAGCGGTGCATTAC 423
QY 407 ATATTGATCATCTTTTGGCTTGATACGTGCATTTAAACAAATACAAAGTCCAAAGTTT 466
DB 424 ACATTGAGCATTTTATCGCTGCTGTTTAGAGTGTGAAAATCCCA-----TAAAAAAGTCT 480
QY 467 TCATTCACATTTTGTGCTGATGTCGAGATATTCGCAACAGCTGAGCTGTTATCTTG 526
DB 481 GTCTGATTTAATCACCAGTGGCGGATGTCGCTCTAAAGCGTTTAACTATTATA 540
QY 527 ACAACTCTCTCAATTTATGCTTCGGAAGATGACGGAATTTGGCTACTATTACGGAC 586
DB 541 AACAAATGC-----AAAATATCTGAATGAAGCATTCAAATTCGTACCAATGAAGCGTC 594
QY 587 GTTATTATCAATGATAGGCAAAAGATGGAGGCTTAAAGATGCTTATAGATGCTTATAGGCA 646
DB 595 GTTTTATGCCATGATAGGATAGCGCTTTGAAGGATGAGCTTGGCTGCTATCA----- 649
QY 647 TTGTTGGAGTATTTGACAAAAAGCCGCTTGATAAGCTGTGCGATGTTGTAGAGAGC 706
DB 650 ----TAGCTTAATGGGCTTAATCACACGCTTTAAGCCCTAGCGAGTATATCCAAAGCC 705
QY 707 GATATCTCAATCTGAGACTGACGAATTTCTGAAACCAATTTGTTTTCGGAGCGGC 766
DB 706 AGTATGATAAAATATCACCGATGAATTTATCATGCCCGCTTGTTTTAAAAAATTTATGCG 765
QY 767 GAGTAAAGATGACGATACCTTATTTTCTTCAATTTATGCTGATGCTGATGCTGCGTCAAA 826
DB 766 GCATGCAAGATGATGAGAGTTTATTTTATCAATTTCAAGATGATAGGCTAGAGAA 825
QY 827 TTTGTGAATTTTGGCTCTCGAACGTTATAAGATCTTAAATAGTTTCGTTCTCCACCCCTA 886
DB 826 TCGTAGCGCTTTAGGCCAAACAAATTCAGTGGCTTTAAGCGCAAGTT-----TTTA 879
QY 887 AAAATATTACAGATTAGTGGATGACCCATACAAATACAAAGATTTCCATTTCCATCGTTAT 946
DB 880 AAAAATCCATATCGCTACCATGACGCTTATGATAACACTTTTCCCTACCTCGTTTAT 939
QY 947 TCCACCTGTGACTCATACTATGTTGCTGCTGAATGGCTTCTCTCAAGAGATTACTC 1006
DB 940 TCCCAAAAGAGCGTTCAAAACACGCTGCTGAAGTGTCTCTCAACAAACCTGACCC 999
QY 1007 AATTTCACTGTGCGAACTGAGAAGTATCTCATGTTTACCTTCTTTTAAATGTTGTC 1066
DB 1000 AAAGCCATATCGCTGAACTGAAAAATACCGCATGTAACCTTTTTCATCATGCGGAG 1059
QY 1067 GAGAAGTTCAATCCAGATGAGAGCGTTGTATGGTTCCGTCACCAAGAAAGTTGCTA 1126

DB 1060 TGGAGACGCGCTTTTAAAAATGAAAAACCGGCTGCTTATCCAAAGCCCTA----AAGTTACCA 1116
QY 1127 CATATGATTTAAACCCAGAAATGAATGCTGCTGGAGTTGCCGAAAAAATGGTCGACAAA 1186
DB 1117 CTTATGACTTAAGCCTGAATGAGCGCTTAAGAGAGTAACCCCTGGCGGTGTAGACAAA 1176
QY 1187 TTGAGTCAGCAGGCGATCCTTTGGTTATGTGCAATTTTGGCTCTCTGATGGTTGGAC 1246
DB 1177 TGAACACTAGCAGCGA-----TTTGATCATTTGTAATTTTCTTAATGCGGATATGTTAGGCG 1233
QY 1247 ATACTGGTAAATTTGAACCTGCGCTCAAGCATGTCAAGCTACTGACGAGGCAATTTGGAA 1306
DB 1234 ATACGGGGAATTTGAAGCGAGCGTCAAGCGGTGAAGCAGTGGATGTATGTTTAGGGG 1293
QY 1307 AGATATTTCAAGCATGCCAAACTTATAATAGTCTTCTTATGTTACTTCCGATCATGGAA 1366
DB 1294 AATCCTTCTCAGTGGCTAAATAATTTGATTAACCCATGCTTTTAAACAGCGATCATGGGA 1353
QY 1367 ATGCTGAGAAGATGA 1381
DB 1354 ATTGGAGCGCATGA 1368

RESULT 12

ABS53606/C

ID ABS53606 standard; DNA; 11769 BP.

XX ABS53606;

XX DT 29-NOV-2002 (first entry)

XX Human Mut/NUDIX DNA sequence.

XX Human; ds; Mut/NUDIX; housecleaning enzyme; cancer;

XX nucleoside triphosphate pyrophosphorylase; cardiovascular disorder;

XX central nervous system disorder; congestive heart failure; ischaemia;

XX arrhythmia; hypertensive vascular disease; peripheral vascular disease;

XX myocardial infarction; angina; brain injury; mood disorder; gene therapy;

XX anxiety; myopathic disorder; neurodegenerative disease; chronic pain;

XX Alzheimer's disease; Parkinson's disease; multiple sclerosis;

XX epilepsy; acquired immunodeficiency syndrome-related pain; migraine;

XX dementia; leukaemia; lymphoma.

XX Homo sapiens.

XX OS WO200266626-A2.

XX PN 29-AUG-2002.

XX PF 14-FEB-2002; 2002WO-EP01541.

XX PR 16-FEB-2001; 2001US-268862P.

XX PA (FARB) BAYER AG.

XX PI Smith TJ;

XX WPI; 2002-674936/72.

XX New human Mut/NUDIX nucleoside triphosphate pyrophosphorylase

XX (Mut/NUDIX NTP) and its encoded protein, useful for identifying

XX modulators of Mut/NUDIX NTP activity, and in gene therapy for treating

XX cancer or heart disease.

XX Disclosure; Fig 3; 135pp; English.

XX The invention relates to an isolated polynucleotide, which encodes a

XX human Mut/NUDIX nucleoside triphosphate pyrophosphorylase, a

XX "housecleaning" enzyme. Also included are Mut/NUDIX expression

XX vectors, host cells, fusion proteins, antibodies and methods of screening

XX for agents that modulate (increase or decrease) the activity of

XX Mut/NUDIX protein or nucleic acid. The Mut/NUDIX NTP polynucleotide and

XX polypeptide are useful for identifying test compounds, that may act as

